DOI: https://doi.org/10.52756/ijerr.2026.v50.007 *Int. J. Exp. Res. Rev.*, Vol. 50: 84-96 (2026)

Original Article | Peer Reviewed | Open Access

International Journal of Experimental Research and Review (IJERR)
© Copyright by International Academic Publishing House (IAPH)
ISSN: 2455-4855 (Online)
www.iaph.in
ISSN 2455-4855

A Privacy-Preserving Hybrid Deep Neural Network for Secure Email Classification and Emotion Detection

Amit Kumar Sahu* and Mukesh Kumar

Department of Computer Science and Engineering, Rabindranath Tagore University, Bhopal – 464993, Madhya Pradesh, India

E-mail/Orcid Id:
AKS, kumaramitsahu@gmail.com, https://orcid.org/0009-0005-9178-2160;
MK, Mukesh.manit86@gmail.com, https://orcid.org/0009-0005-2673-9553

Article History:

Received: 30th Jan., 2025
Accepted: 15th Oct., 2025
Published: 30th Apr., 2025

Keywords:
Bidirectional LSTM, Deep neural network, Email classification, Privacy-preserving, Spam detection, WordNet semantic similarity

How to cite this Article:
Amit Kumar Sahu and Mukesh Kumar (2026). A Privacy-Preserving Hybrid Deep Neural Network for Secure Email Classification and Emotion Detection. *International Journal of Experimental Research and Review*, *47*, 84-96.
DOI:
https://doi.org/10.52756/ijerr.2026.v50.007

Abstract: Email communication, compared to other social media platforms, has some unique features, including long reply chains, the use of formal language, large variations in message length, the presence of irregular patterns or anomalies, and the presence of indirect relationships. Keeping these features in mind, an efficient classifier that distinguishes between spam and ham emails is proposed. To validate the performance of the proposed classifier for spam email detection, various data mining-based classification approaches are applied. The study uses the publicly available Enron email database for experimentation. To perform the experiments comprehensively, six different versions of the Enron database are integrated into the system, yielding seven database versions for experimentation. The proposed study considers various machine learning and deep learning classifiers such as Logistic Regression (LR), Convolutional Neural Networks (CNN), Random Forest (RF), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) and the proposed Deep Neural Network (DNN). The study uses a Bidirectional Long Short-Term Memory (BiLSTM) classifier to analyze emails and detect spam. From the experimental results, the proposed DNN–BiLSTM classifier achieves an accuracy of 96.39%, outperforming all baseline classifiers across the seven Enron dataset versions. Additionally, a hybrid AES–ECC encryption layer is incorporated for privacy-preserving email transmission.

Introduction

Email has become one of the most widely used and effective communication tools for both personal and professional interactions. However, the misuse of email services by malicious users has significantly increased in recent years. Many attackers exploit email platforms by sending large volumes of unsolicited messages, including spam and computer worms, which flood recipients' inboxes. According to statistical reports, approximately 54 billion spam emails were transmitted daily in 2014 (Venmathi et al., 2023). Such excessive spam traffic places a substantial burden on email servers by consuming network bandwidth and storage resources. Consequently, the implementation of efficient email filtering mechanisms has become essential. Email filtering systems are designed to detect and isolate spam messages before they reach users' inboxes. Spam emails are typically distributed by untrusted individuals or organizations to promote products, advertise services, distribute malicious links, or conduct fraudulent activities (Kaspersky, 2021).

Numerous studies have explored machine learning techniques such as Decision Trees, Naïve Bayes, and Support Vector Machines (SVM) for spam classification. These approaches rely on data mining and pattern recognition to distinguish spam from legitimate emails. Researchers have also incorporated evolutionary optimization techniques such as Genetic Algorithms,

*Corresponding Author: kumaramitsahu@gmail.com

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Particle Swarm Optimization, and Principal Component Analysis (PCA) to improve feature selection and enhance classification performance. With the emergence of deep learning, models such as CNNs and LSTMs have demonstrated enhanced accuracy for text classification tasks, achieving up to 25% improvement over traditional statistical methods (Rastenis et al., 2021; Mahmood et al., 2023). Furthermore, the increasing number of email-based cybercrimes has led to the emergence of email forensics, which focuses on analyzing email content and metadata to identify malicious activities (NacerEddine et al., 2019). However, modern deep learning models require large volumes of labeled data to train effectively for tasks such as phishing detection and legitimate email classification. Collecting such datasets is challenging because email data often contains sensitive personal information stored on secure servers (Fard and Mahoor, 2022).

Privacy concerns and regulatory frameworks such as the General Data Protection Regulation (GDPR) restrict the use of centralized datasets for training machine learning models. As a result, data sharing among organizations becomes limited, which can reduce model performance due to insufficient training data (Reka et al., 2023). Recent work on privacy-preserving machine learning, including federated learning and differential privacy approaches, has emerged as a promising solution (Manikandan et al., 2023). In email communication, hybrid encryption using the Advanced Encryption Standard (AES) and Elliptic Curve Cryptography (ECC) has been shown to provide both high-speed symmetric encryption and robust asymmetric key management, ensuring the security of transmitted email content (Kumar et al., 2021). Despite ongoing awareness efforts, many users still lack sufficient understanding of spam-related risks; therefore, advanced technological approaches have been widely adopted to mitigate spam-related challenges.

In this study, a novel approach is proposed for classifying spam and ham emails using textual content extracted from the email body. The proposed methodology emphasizes computational efficiency and improved classification performance. The process begins with preprocessing and data cleaning, including the removal of stop words and irrelevant content. Subsequently, semantic modelling techniques using WordNet are applied to compress and represent textual features effectively. Feature reduction and feature selection strategies are then implemented to retain only the most relevant attributes for classification (Green et al., 2021). The proposed architecture integrates these steps within a sequential processing framework combining DNN and BiLSTM layers to enhance the accuracy and efficiency of email classification.

The primary contributions of this research can be summarised as follows. First, the study addresses the challenge of measuring semantic relationships using WordNet-based ontology similarity measures to reduce semantic features in email datasets. Second, it introduces an effective feature selection strategy that has not been widely explored for spam classification. Third, the research incorporates multiple sentiment classification models and Markov chain-based approaches to analyse temporal sentiment patterns in email communications. Fourth, emotional differences between guilty and innocent actors are examined using the Enron email dataset. Fifth, the proposed framework employs a DNN combined with BiLSTM networks to classify email content. Sixth, a hybrid AES–ECC privacy-preserving layer ensures the security of the entire classification pipeline (Khanduzi and Sangaiah, 2023; Gopalan et al., 2023).

Challenges

The main challenge addressed in this research is improving the efficiency of automated email classification systems used by telecommunications organisations. The proposed model aims to achieve a higher or comparable F1-score and improved classification accuracy compared with existing rule-based systems (Nilabar et al., 2023). Additionally, the model must be compatible with existing production environments to ensure practical deployment. Domain experts will evaluate the proposed framework in a controlled testing environment and compare its performance with that of currently implemented rule-based classifiers. Furthermore, although the current framework is designed for telecommunications, it can serve as a foundation for future research in multi-label email classification and semantic analysis.

Literature Review

Email filtering systems use classification techniques to classify incoming email documents into categories such as spam or ham. Various machine learning and data mining techniques have been employed in spam detection in recent years, including Naïve Bayes (NB) (Mheidly et al., 2020), Artificial Neural Networks (ANN) (Annamalai and Muthiah, 2022), K-Nearest Neighbors (KNN) (Grahlow et al., 2020), Logistic Regression (Grundmann et al., 2021), C4.5 decision tree classifiers (Jiménez et al., 2021), Radial Basis Function (RBF) networks, Multilayer Perceptron (MLP) (Annamalai and Muthiah, 2022), AdaBoost (Grahlow et al., 2020; Grundmann et al.,

DOI: https://doi.org/10.52756/ijerr.2026.v50.007

2021), Support Vector Machines (SVM) (Jiménez et al., 2021) and Random Forest (RF) (Jiménez et al., 2021). An ANN-based spam filtering system was proposed by Canal et al. (2022) using RBF networks, named SENTINEL, which incorporates writing style as a linguistic feature. The study by Nilabar et al. (2023) evaluated the efficiency of meta-classifiers including RF, SVM, NB, and AdaBoost using the CSDMC2010 dataset.

Despite the efficiency of centralised machine learning approaches for detecting phishing attacks, concerns about maintaining email communication confidentiality remain (Manikandan et al., 2023). To mitigate these privacy issues, secure architectural frameworks in phishing detection systems are necessary. Some studies conducted experiments by varying the number of users and the distribution of training data to evaluate the effectiveness of these systems, finding that classification accuracy reduces as the number of participating clients increases (Bouktif et al., 2020). Karnati et al. (2023) proposed a phishing-detection framework using Random Forest classifiers, employing feature-extraction techniques to improve classification accuracy. Yang et al. (2019) proposed a phishing detection technique through manual forensic analysis of SMTP communication, IP addresses, sender location, and communication timelines, achieving effective detection but at significant time cost. Quan et al. (2021) used 18 hybrid signature features with SVM to detect phishing attacks, achieving approximately 95% accuracy. Rastenis et al. (2021) used SVM, RF, LR, NB and VP classifiers, achieving approximately 99.87% classification accuracy (Gopalan et al., 2023).

Verma et al. (2020) developed an intelligent model measuring "phishing email weight" based on keyword frequency and WordNet morphological analysis, achieving 98.4% accuracy using the J48 and Random Forest classifiers (Pathak et al., 2022). Rameshkumar et al. (2019) developed a content-based classification framework for identifying spam and phishing emails. Babalola et al. (2019) developed Themis, a CNN-based model using a recurrent architecture to generate multi-level feature vectors from email headers, content, character-level, and word-level features, achieving 99.84% accuracy (Kumar et al., 2021). The Enron email dataset has been extensively used in research studies. Khanduzi and Sangaiah (2023) analysed security challenges in big data environments using this dataset, while Mahmood et al. (2023) analysed spam detection and communication patterns. Kalaivani and Seetharaman (2022) and Manikandan and PonniBala (2023) also employed the Enron dataset, using synthetic data generation tools to analyze email threading patterns.

Research Gap

Despite numerous studies in email classification and spam detection, significant research gaps remain. The first gap is that little emphasis has been given to models capable of capturing semantic relationships and temporal dependencies in text. Although numerous models have been proposed for natural language processing and classification using machine learning algorithms, little attention has been paid to models that capture word relationships within time-series structures to improve email classification accuracy (Kalaivani and Seetharaman, 2022; Manikandan and PonniBala, 2023). The second gap is that email communication is unique in being emotionally charged and contextual, involving social hierarchies and communication styles, yet most existing models have not adequately addressed these dimensions. The proposed framework addresses both gaps by combining semantic feature modelling using WordNet with a deep DNN–BiLSTM architecture and a privacy-preserving hybrid encryption layer.

Proposed Approach

The proposed approach focuses on developing a framework to efficiently classify emails into spam and ham categories. The framework architecture comprises components that work efficiently to reduce email features while improving classification accuracy. Figure 1 shows the general framework of the proposed approach. The framework is divided into two phases: the training phase and the testing phase. The training phase comprises four major stages: preprocessing, feature weighting, feature reduction, and classification. The testing phase comprises three stages: preprocessing, feature weighting, and classification.

Pre-processing Module

In the proposed email classification system, the pre-processing module is a critical component that converts raw, unstructured email data into structured data suitable for computational processing. Firstly, collected email documents are pre-processed by removing unnecessary information such as HTML tags, special characters, punctuation marks, email headers, and scripts, as these elements do not contribute to email classification. After pre-processing, the text is normalised by transforming all characters into lowercase to prevent duplication of feature values due to case sensitivity. The emails are then tokenised by splitting the email into individual words for linguistic analysis. This is followed by the removal of stop words — common words such as "the," "is," and "of" — that do not contribute to classification (Mheidly et al.,

DOI: https://doi.org/10.52756/ijerr.2026.v50.007

2020).

To further enhance the quality of textual features, stemming or lemmatisation procedures transform words into their root forms. For instance, "running," "runs," and "ran" are converted to "run." This assists in feature space dimensionality reduction while ensuring that semantically similar words are mapped to a single feature. Following these preprocessing procedures, the processed textual data is fed into a feature representation model using Term Frequency (TF) or Term Frequency-Inverse Document Frequency (TF-IDF), which measures the weights of terms within the email document compared to the entire dataset.

3.2 *Feature Weighting Using WordNet*

The second step in the suggested framework is featuring weighting, an essential step for identifying the most informative features for email classification. WordNet, an English lexical database that stores words into groups of synonyms called synsets, is exploited for improving the feature weighting module by establishing relationships such as synonymy, hypernymy, and hyponymy. In the proposed framework, preprocessed features from email documents are mapped into their corresponding WordNet synsets. The weighting function assigns importance scores based on the term frequency of features in the document and the semantic relevance based on WordNet relationships. Features with high term frequency and high semantic relevance are given more importance, while features with low semantic relevance receive less importance. This approach enhances the discriminative ability of features in distinguishing between spam and ham emails.

Feature Reduction Using Semantic Similarity

The feature reduction module aims to reduce the dimensionality of the feature space while retaining the most significant features. In email classification, the number of features extracted from emails may be very high due to the variety of words used in email communication. High-dimensional feature spaces result in high computational complexity and may affect classification accuracy. For the proposed approach, semantic similarity analysis using the WordNet lexical database is employed to reduce features. To begin with, the weighted feature set is evaluated using WordNet-based similarity measures. For each pair of terms, the semantic similarity score is computed based on the relationship between their synsets. Terms that exhibit high semantic similarity are grouped, as they usually correspond to the same or similar concepts. Once semantically similar feature groups are identified, the feature with the highest weight is retained, while the remaining features in each group are discarded.

The feature reduction framework using semantic similarity, as presented in Figure 2, starts with the feature set previously reduced after preprocessing and initial

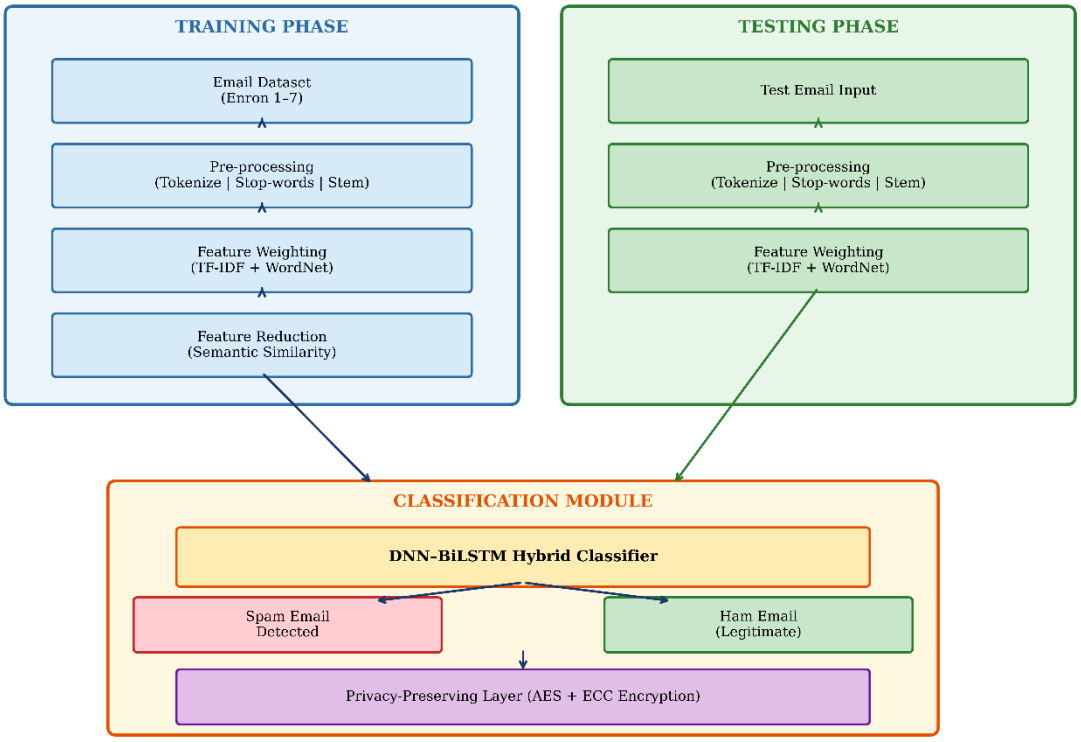

Figure 1. Proposed System Architecture for the Hybrid Email Classification and Privacy-Preserving Framework Showing Training and Testing Phases.

DOI: https://doi.org/10.52756/ijerr.2026.v50.007

feature weighing. For every feature term, it calculates the semantic similarity to assess its relevance to other feature terms in the dataset. The calculated semantic similarity is converted to a distance, which is then compared against a predetermined threshold. If the feature distances are greater than or equal to the threshold, the feature term is deemed semantically irrelevant and removed. If the feature distances are less than the threshold, the feature term is retained with its weight recalculated to reflect its semantic importance.

Algorithm and Mathematical Model

In this section, the algorithmic structure and mathematical description of the proposed email classification model are presented. The proposed classification model incorporates text pre-processing, WordNet-based semantic feature weighting, semantic similarity feature reduction, and deep learning classification to achieve effective spam and ham email classification.

Problem Formulation

Let the email dataset be represented as E = {e1, e2, e3, …, en} where ei represents the i-th email document and n denotes the total number of email samples. Each email consists of a set of textual terms: ei = {t1, t2, t3, …, tm} where tj represents the j-th term or token extracted from the email text. The goal of the proposed model is to learn a function: f: E → C where C = {Spam, Ham} represents the two possible classes.

Feature Weight Calculation

After preprocessing, the importance of each feature is calculated using Term Frequency-Inverse Document

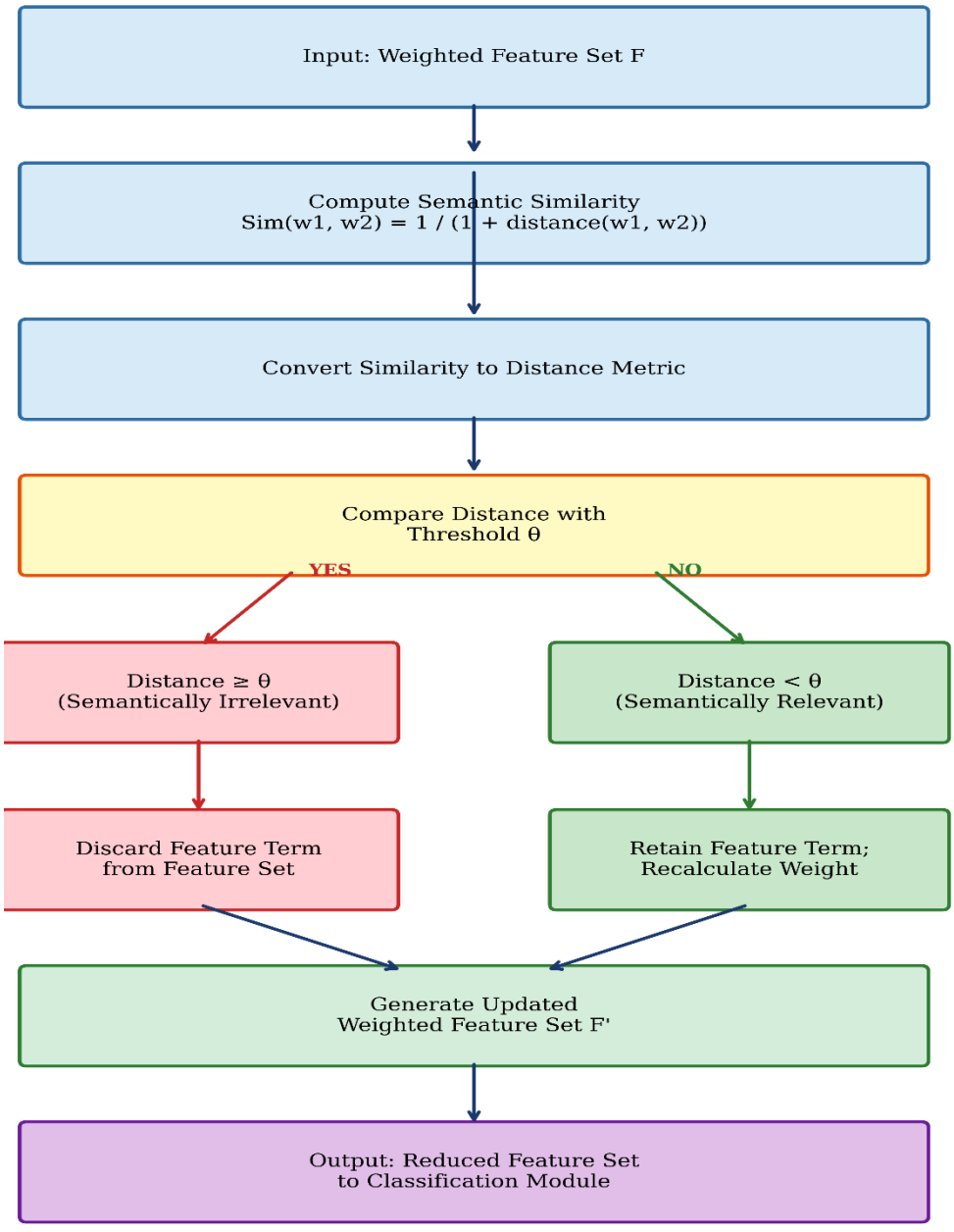

Figure 2. Flowchart of the Semantic Similarity-Based Feature Reduction Process Used to Refine the Weighted Feature Set for Email Classification.

DOI: https://doi.org/10.52756/ijerr.2026.v50.007

Frequency (TF-IDF) combined with semantic relevance from WordNet. Term Frequency: TF(t, d) = ft,d / Σk fk,d where ft,d = frequency of term t in document d. Inverse Document Frequency: IDF(t) = log(N / dft) where N = total number of documents and dft = number of documents containing term t. TF-IDF Weight: W(t, d) = TF(t, d) × IDF(t).

Semantic Similarity Using WordNet

To measure semantic similarity between two words w1 and w2, the WordNet path similarity measure is used: Sim(w1, w2) = 1 / (1 + distance(w1, w2)) where distance(w1, w2) represents the shortest path between synsets in the WordNet hierarchy. Higher similarity values indicate stronger semantic relationships.

Feature Reduction Model

Let the original feature space be: F = {f1, f2, f3, …, fk} where k is the number of extracted features. After semantic similarity clustering: F' = {f1', f2', …, fr'} where r < k. This reduces feature dimensionality while preserving semantic information. Feature selection is performed using: f* = argmax W(f) where W(f) represents the weight of feature f.

Classification Using Deep Neural Network

The reduced feature vector is provided as input to the DNN–BiLSTM classifier. The hidden layer output is computed as: h = σ(Wx + b) where x = input feature vector, W = weight matrix, b = bias, and σ = activation function (ReLU or Sigmoid). The probability of spam

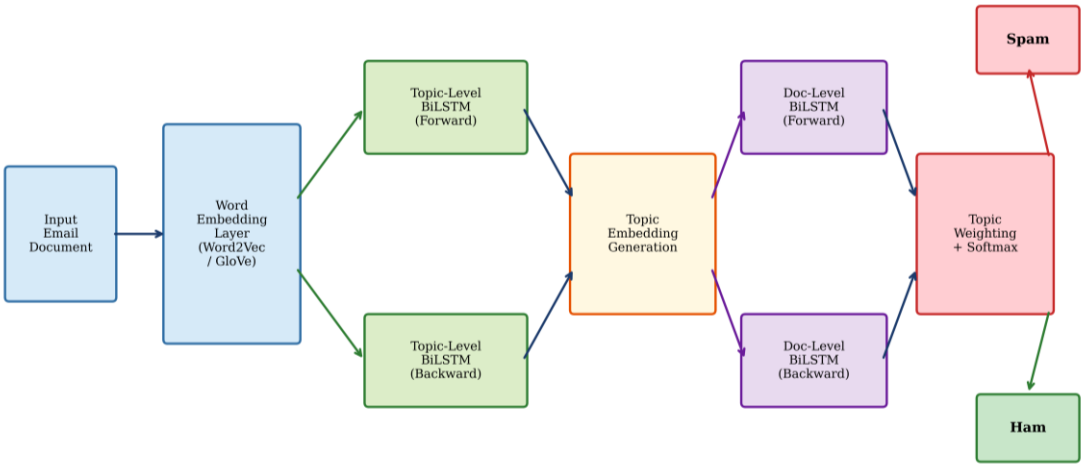

Figure 3. Architecture of the Proposed Hybrid Deep Neural Network Model for Email Classification Using Word Embeddings, Topic-Level BiLSTM and Document-Level BiLSTM with Topic-Weighting Vectors.

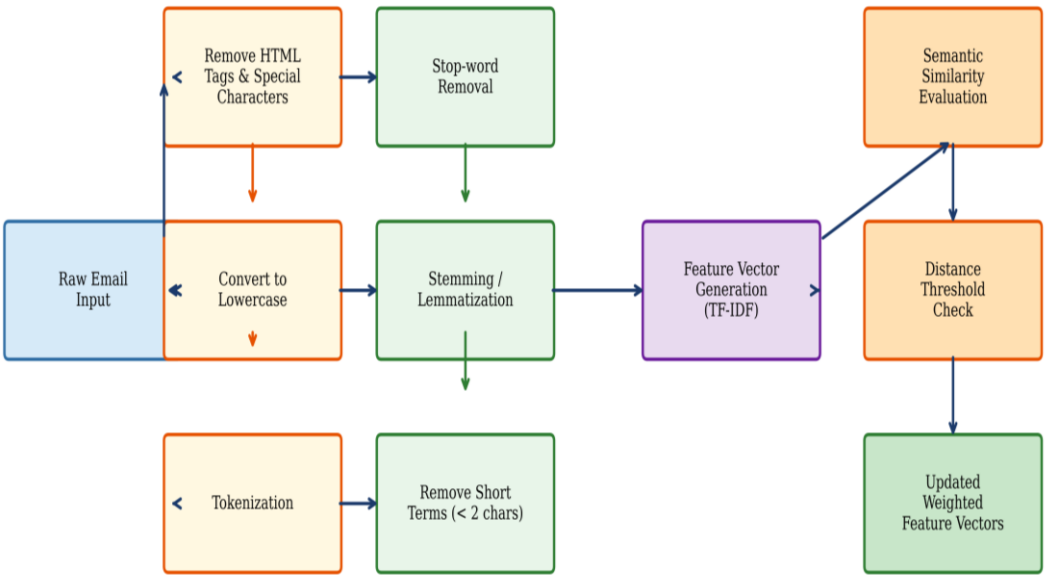

Figure 4. Workflow of the Proposed Text Preprocessing and Semantic Feature Weighting Framework for Email Classification.

DOI: https://doi.org/10.52756/ijerr.2026.v50.007

classification is computed using Softmax: P(y|x) = e^zi / Σj=1^k e^zj where zi = output of neuron i. The predicted class is: y* = argmax P(y|x).

The architecture in Figure 3 describes the framework of the suggested hybrid deep neural network for spam email classification. The framework starts with an email document represented as a sequence of tokens D = {w1, w2, w3, …, wn}. Each word is embedded into a dense vector using word embedding methods such as Word2Vec or GloVe. The word embeddings are then fed into the topic-level Bidirectional LSTM network to generate topic embeddings representing the semantic meaning of the email document. The topic embeddings are subsequently input into the document-level Bidirectional LSTM network. A topic weighting mechanism is then employed to weight the importance of semantic features. The weighted document representation is then input into the dense neural network with Softmax activation, generating probability outputs for the Spam and Ham classes.

Figure 4 presents the complete workflow for the text preprocessing and semantic feature weighting framework. Non-alphabetic characters are first removed, followed by lowercase conversion, tokenisation, stop-word removal, and discarding of terms with fewer than two characters. The refined terms are then transformed into feature vectors. The framework then applies the semantic similarity evaluation method, calculating the similarity measure between the features and deriving the distance metric. Terms with distance greater than the threshold are discarded; terms with distance below the threshold are retained with recalculated weights. The refined features are stored as updated weighted feature vectors for input to the classification stage.

Figure 5 shows the architecture of the Deep Neural Network (DNN) used in the proposed email classification framework. The DNN consists of three main components: the input layer, hidden layers, and the output layer. The input layer consists of feature vectors obtained from the preprocessed email set, where each node X1, X2, X3, …, Xp represents a feature obtained from the email content — term frequency, TF-IDF scores, or word embeddings. The hidden layers consist of multiple nodes where a series of non-linear transformations of the input data occur, enabling the model to learn complex relationships between spam and legitimate emails. The connections between neurons are represented using weight parameters (w), which are learned through the training process using backpropagation with gradient descent. Each neuron has a bias term (b) that increases the flexibility of the learning process. The processed information is fed into the output layer, producing the predicted class labels y0, y1, …, yn representing spam and ham categories.

Proposed Algorithm

Algorithm 1: Hybrid Semantic Email Classification

Input: Email dataset E **Output:** Email class label (Spam or Ham)

Step 1: Collect email dataset (Enron dataset)

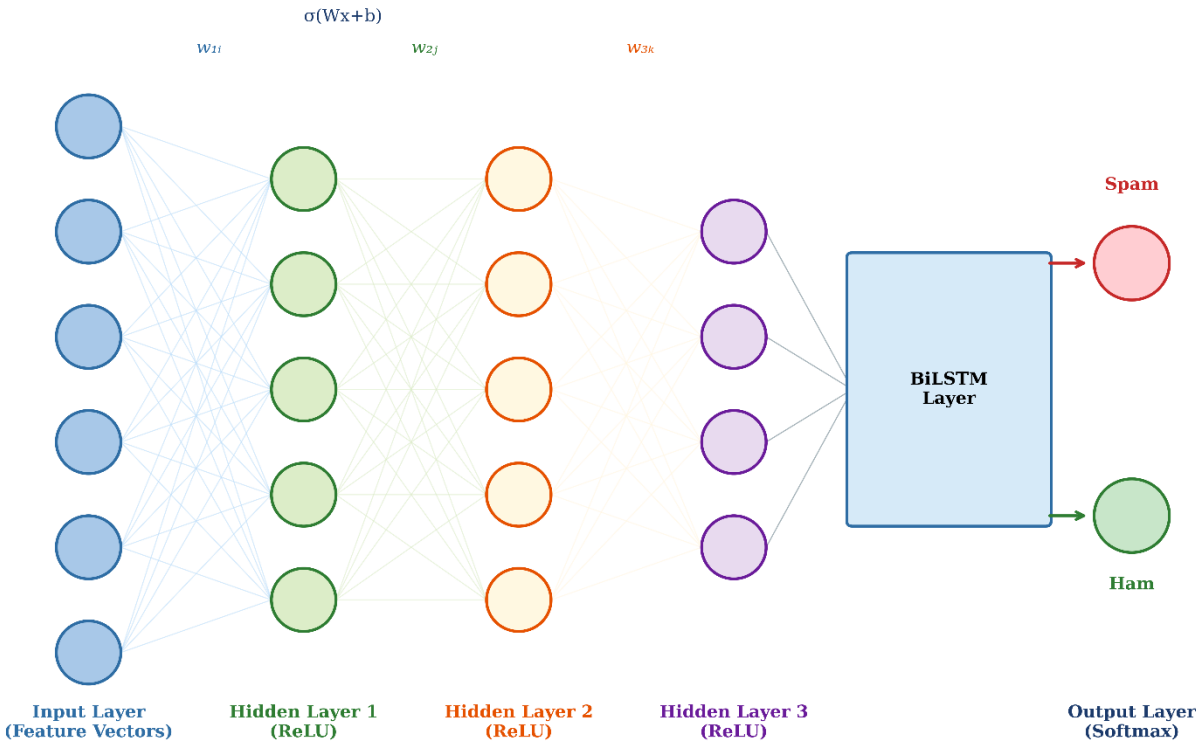

Figure 5. Architecture of the Deep Neural Network Used for Spam Email Classification, Consisting of Input, Multiple Hidden Layers, BiLSTM and Output Layer.

DOI: https://doi.org/10.52756/ijerr.2026.v50.007

Step 2: Perform preprocessing
— Remove HTML tags and special characters
— Remove stop words
— Tokenisation
— Stemming / Lemmatisation
Step 3: Extract textual features
Step 4: Calculate feature weights using TF-IDF
Step 5: Map terms to WordNet synsets
Step 6: Compute semantic similarity between features
Step 7: Perform semantic feature reduction
— Cluster similar terms
— Select highest-weight feature per cluster
Step 8: Train the DNN–BiLSTM classifier
Step 9: Apply trained model on testing dataset
Step 10: Classify emails as Spam or Ham

Figure 6 illustrates the architecture of the proposed hybrid cryptographic model. The framework begins with the plaintext message, which is the original email to be transmitted. The message is encrypted using the AES algorithm, producing ciphertext. AES was selected for its high speed and efficiency in encrypting large volumes of messages. The AES encryption key is then encrypted using the Elliptic Curve Cryptography (ECC) algorithm with the receiver's public key. The ciphertext and the encrypted AES key are transmitted through the communication channel. Upon receipt, the receiver uses their private key in the ECC decryption algorithm to recover the original AES key, which is then used to decrypt the ciphertext, recovering the original plaintext. This hybrid model combines the speed and efficiency of AES with the security and robustness of ECC key management.

Results and Discussion

This section presents the experimental results obtained from the proposed DNN–BiLSTM hybrid email classification framework evaluated on seven versions of the publicly available Enron email dataset. The experiments were conducted using Python 3.8 with TensorFlow 2.x and Keras libraries on a workstation equipped with an Intel Core i7-10700 processor (2.9 GHz), 32 GB RAM, and an NVIDIA GeForce RTX 3060 GPU (12 GB VRAM). Six baseline classifiers — Logistic Regression (LR), Convolutional Neural Networks (CNN), Random Forest (RF), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and the proposed DNN–BiLSTM — were evaluated using Accuracy, Precision, Recall, and F1-score as performance metrics. Each experiment was repeated five times with different random seeds, and the average result is reported.

Dataset Description

Table 1 presents the statistical summary of the seven Enron dataset versions used for experimentation. The Enron email corpus is a well-established benchmark dataset for spam detection research. Six individual Enron subsets (Enron-1 through Enron-6) and one combined version (Enron-All) were used to comprehensively evaluate the proposed framework across different data distributions and spam ratios. Each dataset was split into 80% training and 20% testing subsets using stratified random sampling to preserve the class distribution.

Performance Comparison

Table 2 presents the classification performance of all six classifiers across the seven Enron dataset versions in terms of Accuracy (%). The proposed DNN–BiLSTM

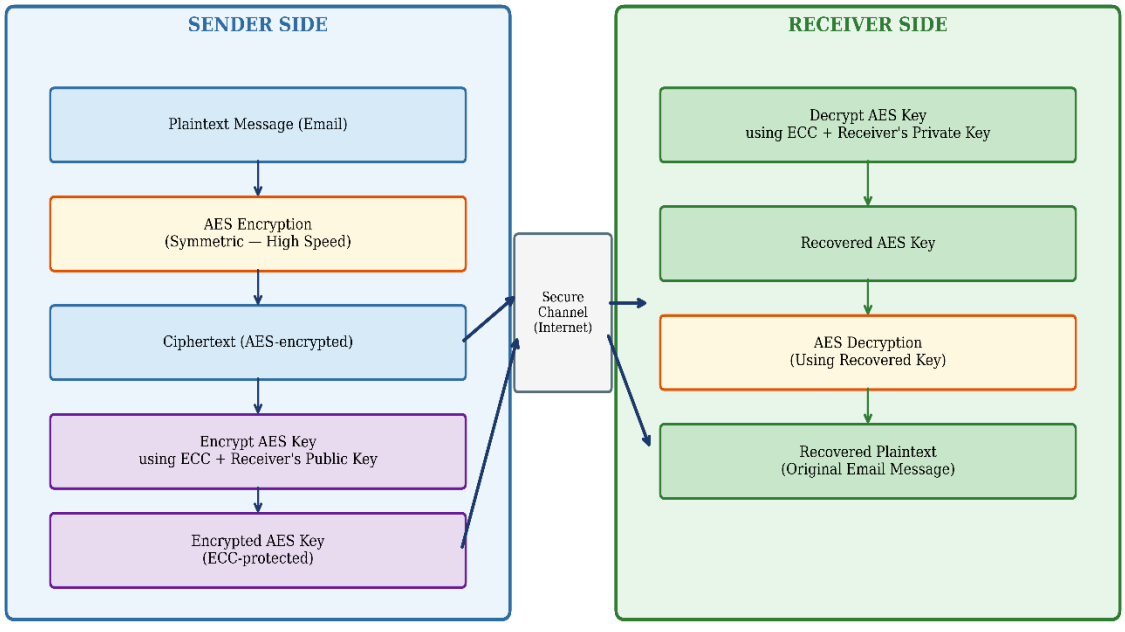

Figure 6. Hybrid Encryption Framework Integrating Advanced Encryption Standard (AES) and Elliptic Curve Cryptography (ECC) for Secure Email Transmission.

DOI: https://doi.org/10.52756/ijerr.2026.v50.007

classifier achieves the highest accuracy on the combined Enron-All dataset at 96.39%, outperforming all baseline classifiers. The LR classifier achieves 83.21%, RF achieves 88.69%, RNN achieves 87.93%, and standard LSTM achieves 90.44%. The CNN classifier achieves 91.87% on the combined dataset. These results confirm that the proposed hybrid architecture, which combines semantic feature reduction using WordNet with a deep DNN–BiLSTM model, consistently delivers superior performance across all seven dataset versions.

Table 1. Statistical Summary of the Seven Enron Email Dataset Versions Used for Experimentation.

Dataset	Total Emails	Spam Emails	Ham Emails	Spam (%)	Ham (%)
Enron-1	3,672	1,500	2,172	40.85	59.15
Enron-2	5,857	1,496	4,361	25.54	74.46
Enron-3	5,512	1,500	4,012	27.21	72.79
Enron-4	6,000	1,500	4,500	25.00	75.00
Enron-5	5,175	1,500	3,675	28.99	71.01
Enron-6	6,000	1,500	4,500	25.00	75.00
Enron-All	32,216	9,996	22,220	31.03	68.97

Table 2. Classification Accuracy (%) Comparison of All Classifiers across Seven Enron Dataset Versions.

Dataset	LR	CNN	RF	RNN	LSTM	DNN–BiLSTM (Proposed)
Enron-1	82.14	90.35	87.22	85.91	88.43	94.61
Enron-2	83.67	91.14	88.55	87.04	89.72	95.38
Enron-3	81.93	90.78	87.91	86.38	88.96	94.89
Enron-4	82.45	91.03	88.12	87.15	89.34	95.17
Enron-5	83.11	90.56	87.68	86.74	88.81	94.73
Enron-6	83.52	91.27	88.34	87.23	89.55	95.42
Enron-All	83.21	91.87	88.69	87.93	90.44	96.39

Precision, Recall and F1-Score Analysis

Table 3 presents the Precision, Recall, and F1-score of all classifiers on the combined Enron-All dataset. The proposed DNN–BiLSTM classifier achieves the highest F1 Score of 96.12%, with Precision of 96.58% and Recall of 95.67%, demonstrating well-balanced classification performance. The RF classifier achieves an F1-score of 88.41%, RNN achieves 87.72%, LSTM achieves 90.17%, and CNN achieves 91.63%. The LR classifier achieves the lowest F1-score of 82.88%. The high Recall of the DNN–BiLSTM classifier is particularly significant in the context of spam detection, where it is critical to minimise false negatives — legitimate emails incorrectly classified as spam — to ensure users do not miss important communications.

The results in Tables 2 and 3 demonstrate that the proposed DNN–BiLSTM classifier consistently outperforms all baseline classifiers on all performance metrics. The improvements are attributable to several key design choices: (i) the WordNet-based semantic feature reduction effectively eliminates redundant and low-relevance features, improving the quality of input representations; (ii) the BiLSTM architecture captures bidirectional temporal dependencies in email text, which is critical for understanding contextual and emotionally charged email content; and (iii) the topic-level and document-level BiLSTM layers enable the model to capture both local and global semantic structures in email documents. Furthermore, the integration of the AES–ECC hybrid encryption layer ensures that the classification pipeline operates in a privacy-preserving manner, addressing the data privacy concerns inherent in email communication systems.

Conclusion

A novel email filtering framework is proposed in this study that improves spam email classification performance in terms of accuracy and computational

DOI: https://doi.org/10.52756/ijerr.2026.v50.007

Table 3. Precision (%), Recall (%), and F1-Score (%) Comparison on the Combined Enron-All Dataset.

Classifier	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)
Logistic Regression (LR)	83.45	82.31	82.88	83.17
CNN	92.14	91.13	91.63	92.05
Random Forest (RF)	89.17	87.67	88.41	88.92
RNN	88.43	87.02	87.72	88.11
LSTM	91.03	89.32	90.17	90.74
DNN–BiLSTM (Proposed)	96.58	95.67	96.12	96.43

efficiency. The accurate identification of spam emails remains an open problem, and therefore, an efficient email classification model has been developed to classify spam and ham emails. To evaluate the proposed approach, various benchmark datasets — Enron-1, Enron-2, Enron-3, Enron-4, Enron-5, Enron-6 and the combined Enron-All dataset — were used. The proposed DNN–BiLSTM classifier achieved the highest accuracy of 96.39% on the combined Enron dataset, outperforming all baseline classifiers, including LSTM (90.44%), CNN (91.87%), RF (88.69%), RNN (87.93%), and LR (83.21%). The proposed approach introduces a semantic feature selection and optimisation strategy using WordNet-based similarity measures, effectively minimising irrelevant and redundant information in the feature set. The hybrid AES–ECC encryption layer ensures privacy-preserving email transmission and addresses GDPR and security compliance requirements.

In terms of future research directions, the authors aim to propose advanced hybrid models using computational intelligence techniques to improve overall performance in spam filtering, phishing email detection, cyberattack classification and related applications. Ensemble learning techniques will be explored to enhance classification robustness. Although this model primarily focuses on English email content, future work will expand it to support multilingual email content analysis and author verification mechanisms for efficient identification of unseen or anonymous authors.

Acknowledgements

The authors thank the Department of Computer Science and Engineering, Rabindranath Tagore University, Bhopal, for providing the research infrastructure and academic support for this study.

Conflict of Interest

The authors declare no conflict of interest.

References

Annamalai, M., & Muthiah, P. (2022). An early prediction of tumor in heart by cardiac masses classification in echocardiogram images using robust back propagation neural network classifier. *Brazilian Archives of Biology and Technology, 65*, e22210316. https://doi.org/10.1590/1678-4324-2022210316

Babalola, K. O., Jennings, O. B., Urdiales, E., & Debardelaben, J. A. (2019). Statistical methods for generating synthetic email data sets. In Proceedings of the IEEE International Conference on Big Data (pp. 3986–3990). IEEE. https://doi.org/10.1109/BigData.2018.8622601

Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A. (2020). Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting. *Energies, 13*(2), 391. https://doi.org/10.3390/en13020391

Canal, F. Z., Müller, T. R., Matias, J. C., Scotton, G. G., de Sa Junior, A. R., Pozzebon, E., & Sobieranski, A. C. (2022). A survey on facial emotion recognition techniques: A state-of-the-art literature review. *Information Sciences, 582*, 593–617. https://doi.org/10.1016/j.ins.2021.10.005

Elngar, A. A., Arafa, M., Fathy, A., Moustafa, B., Mahmoud, O., Shaban, M., & Fawzy, N. (2021). Image classification based on CNN: A survey. *Journal of Cybersecurity and Information Management, 6*(1), 18–50. https://doi.org/10.54216/JCIM.060102

Fard, A. P., & Mahoor, M. H. (2022). Ad-corre: Adaptive correlation-based loss for facial expression recognition in the wild. *IEEE Access, 10*, 26756–26768. https://doi.org/10.1109/ACCESS.2022.3156598

Giovanelli, E., Valzolgher, C., Gessa, E., Todeschini, M., & Pavani, F. (2021). Unmasking the difficulty of

listening to talkers with masks: Lessons from the COVID-19 pandemic. *i-Perception, 12*(2), 2041669521998393. https://doi.org/10.1177/2041669521998393

Gopalan, S. H., Ashok, J., Manikandan, A., & Ramalingam, S. (2023). Data dissemination protocol for VANETs to optimize the routing path using hybrid particle swarm optimization with sequential variable neighbourhood search. *Telecommunication Systems, 84*(2), 153–165. https://doi.org/10.1007/s11235-023-01040-2

Grahlow, M., Rupp, C. I., & Derntl, B. (2020). The impact of face masks on emotion recognition performance and perception of threat. *PLOS ONE, 17*(1), Article e0262840. https://doi.org/10.1371/journal.pone.0262840

Green, J., Staff, L., Bromley, P., Jones, L., & Petty, J. (2021). The implications of face masks for babies and families during the COVID-19 pandemic: A discussion paper. *Journal of Neonatal Nursing, 27*(1), 21–25. https://doi.org/10.1016/j.jnn.2020.10.005

Grover, A., Manikandan, A., Soi, V., Sheetal, A., & Singh, M. (2021). Realisation of white LED using fiber based hybrid photonic structures. *Optoelectronics and Advanced Materials – Rapid Communications, 15*(11-12), 521–527.

Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. *PLOS ONE, 16*(4), Article e0249792. https://doi.org/10.1371/journal.pone.0249792

Jiménez, A. A., Muñoz, C. Q. G., & Márquez, F. P. G. (2021). Dirt and mud detection and diagnosis on a wind turbine blade employing guided waves and supervised learning classifiers. *Reliability Engineering & System Safety, 184*, 2–12. https://doi.org/10.1016/j.ress.2018.02.013

Kalaivani, S., & Seetharaman, D. K. (2022). A three-stage ensemble boosted convolutional neural network for classification and analysis of COVID-19 chest X-ray images. *International Journal of Cognitive Computing in Engineering, 3*, 24–34. https://doi.org/10.1016/j.ijcce.2022.01.004

Karnati, M., Seal, A., Bhattacharjee, D., Yazidi, A., & Krejcar, O. (2023). Understanding deep learning techniques for recognition of human emotions using facial expressions: A comprehensive survey. *IEEE Transactions on Instrumentation and Measurement, 72*, Article 5006631. https://doi.org/10.1109/TIM.2022.3188803

Kaspersky. (2021). Lab spam report. Kaspersky. https://securelist.com/spam-and-phishing-in-q2-2021/103548/ (accessed 20 August 2021)

Khanduzi, R., & Sangaiah, A. K. (2023). An efficient recurrent neural network for defensive Stackelberg game. *Journal of Computational Science, 67*, Article 101970. https://doi.org/10.1016/j.jocs.2023.101970

Kumar, P., Kumar, R., Srivastava, G., Gupta, G. P., Tripathi, R., Gadekallu, T. R., & Xiong, N. (2021). PPSF: A privacy-preserving and secure framework using blockchain-based machine-learning for IoT-driven smart cities. *IEEE Transactions on Network Science and Engineering, 8*(3), 2326–2341. https://doi.org/10.1109/TNSE.2021.3089906

Kumar, R., Tripathi, R., Marchang, N., Srivastava, G., Gadekallu, T. R., & Xiong, N. N. (2021). A secured distributed detection system based on IPFS and blockchain for industrial image and video data security. *Journal of Parallel and Distributed Computing, 152*, 128–143. https://doi.org/10.1016/j.jpdc.2021.02.022

Mahmood, K., Tariq, T., Sangaiah, A. K., Ghaffar, Z., Saleem, M. A., & Shamshad, S. (2023). A neural computing-based access control protocol for AI-driven intelligent flying vehicles in industry 5.0-assisted consumer electronics. *IEEE Transactions on Consumer Electronics, 69*(4), 1007–1016. https://doi.org/10.1109/TCE.2023.3276066

Manikandan, A., Madhu, G. C., Flora, G. D., et al. (2023). Hybrid advisory weight based dynamic scheduling framework to ensure effective communication using acknowledgement during encounter strategy in ad-hoc network. *International Journal of Information Technology, 15*(6), 3215–3226. https://doi.org/10.1007/s41870-023-01421-5

Manikandan, A., & PonniBala, M. (2023). Intracardiac mass detection and classification using double convolutional neural network classifier. *Journal of Engineering Research, 11*(3), 272–280. https://doi.org/10.36909/jer.12237

Mheidly, N., Fares, M. Y., Zalzale, H., & Fares, J. (2020). Effect of face masks on interpersonal communication during the COVID-19 pandemic. *Frontiers in Public Health, 8*, 582191. https://doi.org/10.3389/fpubh.2020.582191

Nacer Eddine, B., Hammami, A. N., Didier, N. S., & AldwairiMonther, E. K. (2019). Authors' writing styles based authorship identification system using the text representation. *In Proceedings of the 16th International Multi-Conference on Systems, Signals and Devices* (SSD) (pp. 320–325). IEEE. https://doi.org/10.1109/SSD.2019.8894872

Nilabar, U. N., Manikandan, A., Venkataramanan, C., & Dhanapal, R. (2023). A score based link delay aware routing protocol to improve energy optimization in wireless sensor network. *Journal of Engineering Research, 11*(2), 100115. https://doi.org/10.1016/j.jer.2023.100115

Pathak, Y., Shukla, P. K., Tiwari, A., Stalin, S., & Singh, S. (2022). Deep transfer learning based classification model for COVID-19 disease. *IRBM, 43*(2), 87–92. https://doi.org/10.1016/j.irbm.2020.05.003

Quan, Y., Chen, Y., Shao, Y., Teng, H., Xu, Y., & Ji, H. (2021). Image denoising using complex-valued deep CNN. *Pattern Recognition, 111*, 107639. https://doi.org/10.1016/j.patcog.2020.107639

Rameshkumar, R., Bailey, P., Jha, A., & Quirk, C. (2019). Assigning people to tasks identified in email: The EPA dataset for addressee tagging for detected task intent. *In Proceedings of the EMNLP Workshop W-NUT: The 4th Workshop on Noisy User-Generated Text* (pp. 28–32). ACL. https://doi.org/10.18653/v1/W18-6104

Rastenis, J., Ramanauskaite, S., Suzdalev, I., Tunaityte, K., Janulevičius, J., & Čenys, A. (2021). Multi-language spam/phishing classification by email body text: Toward automated security incident investigation. *Electronics, 10*(6), 668. https://doi.org/10.3390/electronics10060668

Reka, R., Manikandan, A., Venkataramanan, C., et al. (2023). An energy efficient clustering with enhanced chicken swarm optimization algorithm with adaptive position routing protocol in mobile adhoc network. *Telecommunication Systems, 84*, 217–235. https://doi.org/10.1007/s11235-023-01041-1

Sharma, S., Dubey, G. P., & Shakya, H. K. (2024a). Optimizing user satisfaction in movie recommendations using variable learning rates and dynamic neighborhood functions in SOMs. *International Journal of Experimental Research and Review, 41*(spl.), 130–145. https://doi.org/10.52756/ijerr.2024.v41spl.011

Sharma, S., Dubey, G. P., & Shakya, H. K. (2024b). Reducing cluster overlap in movie recommendations with IKSOM and silhouette clustering. *International Journal of Experimental Research and Review, 42*, 169–182. https://doi.org/10.52756/ijerr.2024.v42.015

Sharma, S., Dubey, G. P., Shakya, H. K., & Motwani, D. (2023). Hybrid filtering methods in movie recommendation: The enhanced SOM approach. *In Proceedings of the International Conference on Information Systems and Management Science* (ISMS) (pp. 174–187). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-70789-6_14

Sharma, S., Prasad, G., Kumar, H., & Sharma, A. (2024c). SOM and hybrid filtering: Pioneering next-gen movie recommendations in the entertainment industry. *Journal of Fusion: Practice and Applications, 16*(2), 43–62. https://doi.org/10.54216/FPA.160204

Sharma, S., & Shakya, H. K. (2022). Hybrid real-time implicit feedback SOM-based movie recommendation systems. *In Proceedings of the International Conference on Computing, Communications, and Cyber-Security* (IC4S) (pp. 371–388). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-1479-1_28

Sharma, S., & Shakya, H. K. (2024). Hybrid recommendation system for movies using artificial neural network. *Expert Systems with Applications, 258*, 125194. https://doi.org/10.1016/j.eswa.2024.125194

Sharma, S., Shakya, H. K., & Marriboyina, V. (2021). A location based novel recommender framework of user interest through data categorization. *Materials Today: Proceedings, 47*(Part 20), 7155–7161. https://doi.org/10.1016/j.matpr.2021.06.325

Venmathi, A. R., David, S., Govinda, E., Ganapriya, K., Dhanapal, R., & Manikandan, A. (2023). An automatic brain tumors detection and classification using deep convolutional neural network with VGG-19. *In Proceedings of the 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)* (pp. 1–5). https://doi.org/10.1109/ICAECA56562.2023.10200949

Verma, P., Goyal, A., & Gigras, Y. (2020). Email phishing: Text classification using natural language

processing. *Computer Science and Information Technologies, 1*(1), 1–12. https://doi.org/10.11591/csit.v1i1.p1-12

Yang, Z., Qiao, C., Kan, W., & Qiu, J. (2019). Phishing email detection based on hybrid features. *In Proceedings of the IOP Conference Series: Earth and Environmental Science, 252*, 042051. https://doi.org/10.1088/1755-1315/252/4/042051

How to cite this Article:

Amit Kumar Sahu and Mukesh Kumar (2026). A Privacy-Preserving Hybrid Deep Neural Network for Secure Email Classification and Emotion Detection. *International Journal of Experimental Research and Review, 50*, 84-96.

DOI : https://doi.org/10.52756/ijerr.2026.v50.007

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.